

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number:

10/525,020

Source:

PCT

Date Processed by STIC:

3-4-05

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 03/04/2005

PATENT APPLICATION: US/10/525,020

TIME: 12:50:18

Input Set : A:\PH-1716 Sequence.TXT

Output Set: N:\CRF4\03042005\J525020.raw

3 <110> APPLICANT: YAMAMOTO, Kazuo
 4 Sumitomo Corporation
 6 <120> TITLE OF INVENTION: Carbohydrate library constructed by Gene Alteration of Cargo
 7 Receptors
 9 <130> FILE REFERENCE: PH-1716PCT
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/525,020
 C--> 12 <141> CURRENT FILING DATE: 2005-02-18
 15 <150> PRIOR APPLICATION NUMBER: JP 2002-238559
 16 <151> PRIOR FILING DATE: 2002-08-19
 18 <160> NUMBER OF SEQ ID NOS: 17
 20 <170> SOFTWARE: PatentIn Ver. 2.0
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 2768
 24 <212> TYPE: DNA
 25 <213> ORGANISM: Homo sapiens
 27 <220> FEATURE:
 28 <221> NAME/KEY: CDS
 29 <222> LOCATION: (22)..(1554)
 31 <220> FEATURE:
 32 <221> NAME/KEY: sig_peptide
 33 <222> LOCATION: (22)..(112)
 35 <400> SEQUENCE: 1
 36 ggtcgcggttc cagaatccaa g atg gcg gga tcc agg caa agg ggt ctc cgg 51
 37 Met Ala Gly Ser Arg Gln Arg Gly Leu Arg
 38 1 5 10
 40 gcc aga gtt cgg ccg ctg ttc tgc gcc ttg ctg tca ctc ggt cgc 99
 41 Ala Arg Val Arg Pro Leu Phe Cys Ala Leu Leu Ser Leu Gly Arg
 42 15 20 25
 44 ttc gtc cgg ggc gac ggc gtg gga gga gac ccc gcg gtc gcg ttg cca 147
 45 Phe Val Arg Gly Asp Gly Val Gly Gly Asp Pro Ala Val Ala Leu Pro
 46 30 35 40
 48 cat cgc cgt ttc gag tac aaa tac agc ttc aag ggg ccg cac ctg gtg 195
 49 His Arg Arg Phe Glu Tyr Lys Tyr Ser Phe Lys Gly Pro His Leu Val
 50 45 50 55
 52 cag agc gac ggg acc gtg ccc ttc tgg gcc cac gcg ggg aat gct att 243
 53 Gln Ser Asp Gly Thr Val Pro Phe Trp Ala His Ala Gly Asn Ala Ile
 54 60 65 70
 56 cca agt tca gat caa att cga gta gca cca tct tta aaa agc caa aga 291
 57 Pro Ser Ser Asp Gln Ile Arg Val Ala Pro Ser Leu Lys Ser Gln Arg
 58 75 80 85 90
 60 ggc tca gtg tgg aca aag aca aaa gcg gcc ttt gag aac tgg gaa gtt 339
 61 Gly Ser Val Trp Thr Lys Thr Lys Ala Ala Phe Glu Asn Trp Glu Val
 62 95 100 105

(pg. 6)

RAW SEQUENCE LISTING

DATE: 03/04/2005

PATENT APPLICATION: US/10/525,020

TIME: 12:50:18

Input Set : A:\PH-1716 Sequence.TXT

Output Set: N:\CRF4\03042005\J525020.raw

```

64 gag gtg aca ttt cga gtg act gga aga ggt cga att gga gct gat ggc 387
65 Glu Val Thr Phe Arg Val Thr Gly Arg Gly Arg Ile Gly Ala Asp Gly
66      110      115      120
68 cta gca att tgg tat gca gaa aat caa ggc ttg gag ggc cct gtg ttt 435
69 Leu Ala Ile Trp Tyr Ala Glu Asn Gln Gly Leu Glu Gly Pro Val Phe
70      125      130      135
72 gga tca gct gat ctg tgg aat ggt gtt gga ata ttt ttt gat act ttt 483
73 Gly Ser Ala Asp Leu Trp Asn Gly Val Gly Ile Phe Phe Asp Thr Phe
74      140      145      150
76 gac aat gat gga aag aaa aat aat cct gct ata gta att ata ggc aac 531
77 Asp Asn Asp Gly Lys Lys Asn Asn Pro Ala Ile Val Ile Ile Gly Asn
78 155      160      165      170
80 aat gga caa atc cat tat gac cat caa aat gac ggg gct agt caa gct 579
81 Asn Gly Gln Ile His Tyr Asp His Gln Asn Asp Gly Ala Ser Gln Ala
82      175      180      185
84 ttg gca agt tgc cag agg gac ttc cgc aac aaa ccc tat cct gtc cga 627
85 Leu Ala Ser Cys Gln Arg Asp Phe Arg Asn Lys Pro Tyr Pro Val Arg
86      190      195      200
88 gca aag att acc tat tac cag aac aca ctg aca gta atg atc aat aat 675
89 Ala Lys Ile Thr Tyr Tyr Gln Asn Thr Leu Thr Val Met Ile Asn Asn
90      205      210      215
92 ggc ttt aca cca gat aaa aat gat tat gaa ttt tgt gcc aaa gtg gaa 723
93 Gly Phe Thr Pro Asp Lys Asn Asp Tyr Glu Phe Cys Ala Lys Val Glu
94      220      225      230
96 aat atg att atc cct gca caa ggg cat ttt gga ata tct gct gca act 771
97 Asn Met Ile Ile Pro Ala Gln Gly His Phe Gly Ile Ser Ala Ala Thr
98 235      240      245      250
100 gga ggt ctt gca gat gac cat gat gtc ctt tct ttt ctg act ttc cag 819
101 Gly Gly Leu Ala Asp Asp His Asp Val Leu Ser Phe Leu Thr Phe Gln
102      255      260      265
104 ttg act gaa cct gga aaa gag ccg ccc aca cca gat aaa gaa att tcg 867
105 Leu Thr Glu Pro Gly Lys Glu Pro Pro Thr Pro Asp Lys Glu Ile Ser
106      270      275      280
108 gaa aag gaa aaa gaa aag tat cag gag gaa ttt gag cac ttt caa caa 915
109 Glu Lys Glu Lys Glu Lys Tyr Gln Glu Glu Phe Glu His Phe Gln Gln
110      285      290      295
112 gaa ttg gat aaa aaa aaa gag gaa ttc cag aag ggc cac ccc gac ctc 963
113 Glu Leu Asp Lys Lys Lys Glu Glu Phe Gln Lys Gly His Pro Asp Leu
114      300      305      310
116 caa ggg cag cct gcg gag gaa ata ttt gag agt gta gga gat cga gag 1011
117 Gln Gly Gln Pro Ala Glu Glu Ile Phe Glu Ser Val Gly Asp Arg Glu
118 315      320      325      330
120 cta aga caa gtc ttt gaa gga cag aat cgt att cat ctt gaa atc aag 1059
121 Leu Arg Gln Val Phe Glu Gly Gln Asn Arg Ile His Leu Glu Ile Lys
122      335      340      345
124 cag ctg aac cgg cag tta gat atg att ctt gat gaa cag aga aga tat 1107
125 Gln Leu Asn Arg Gln Leu Asp Met Ile Leu Asp Glu Gln Arg Arg Tyr
126      350      355      360
128 gtc tct tcc tta aca gag gaa atc tct aaa aga gga gca gga atg cct 1155

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/525,020

DATE: 03/04/2005

TIME: 12:50:18

Input Set : A:\PH-1716 Sequence.TXT

Output Set: N:\CRF4\03042005\J525020.raw

```

129 Val Ser Ser Leu Thr Glu Glu Ile Ser Lys Arg Gly Ala Gly Met Pro
130      365      370      375
132 ggg cag cat ggg cag att act caa caa gaa ctg gat act gtt gtg aaa 1203
133 Gly Gln His Gly Gln Ile Thr Gln Gln Glu Leu Asp Thr Val Val Lys
134      380      385      390
136 act cag cat gag att ctg aga caa gta aat gaa atg aaa aat tcc atg 1251
137 Thr Gln His Glu Ile Leu Arg Gln Val Asn Glu Met Lys Asn Ser Met
138 395      400      405      410
140 agt gaa acc gtc aga ctg gtc agt gga atg cag cac cct ggc tct gct 1299
141 Ser Glu Thr Val Arg Leu Val Ser Gly Met Gln His Pro Gly Ser Ala
142      415      420      425
144 gga ggc gtc tat gag aca aca cag cac ttc att gac atc aaa gag cac 1347
145 Gly Gly Val Tyr Glu Thr Thr Gln His Phe Ile Asp Ile Lys Glu His
146      430      435      440
148 ctg cac ata gta aag agg gac ata gat aac tta gtg cag cga aat atg 1395
149 Leu His Ile Val Lys Arg Asp Ile Asp Asn Leu Val Gln Arg Asn Met
150      445      450      455
152 cca tca aat gaa aag ccg aaa tgc cca gaa cta cca cca ttt cca tca 1443
153 Pro Ser Asn Glu Lys Pro Lys Cys Pro Glu Leu Pro Pro Phe Pro Ser
154      460      465      470
156 tgt ttg tct acg gtc cac ttc att ata ttt gtt gtg gtg caa act gta 1491
157 Cys Leu Ser Thr Val His Phe Ile Ile Phe Val Val Val Gln Thr Val
158 475      480      485      490
160 tta ttc att ggt tat atc atg tat agg tct cag caa gaa gca gct gcc 1539
161 Leu Phe Ile Gly Tyr Ile Met Tyr Arg Ser Gln Gln Glu Ala Ala Ala
162      495      500      505
164 aaa aaa ttc ttt tga ctaccatttt cctgtgtact tcatctatatt gtgtacaaaa 1594
165 Lys Lys Phe Phe
166      510
168 tgagtcgttt tgagggaatt taagtattta aattgcttca tagtctaaat tattaatttt 1654
170 cttaataaaa taactgttta aacattgatt tgcagtttaag aataaacctt aaagcaaaga 1714
172 caaccacatt ttaatttggt cacagtagt aaactgtct aaatttcagt gaatttctgg 1774
174 tcagtatgat gcagcctctg agcagaatat tgaccagtaa gagggtaaat aaagtggggg 1834
176 caaccctgga tatgaatggt acccctaag tctccaatat tgcaggtttc cctgtataac 1894
178 gtaaacacac ttgccctcat gcctcccaga atatgaggtc taattaagaa gtccatcagg 1954
180 tttattttgt aaccaaagtc ttttttagag gtcagacttc ctaatcaaag gcctgggcct 2014
182 gcagtcctt tcatcttaat gcaacttctt ttgaaatcaa agaataattt gtctgagagc 2074
184 ttttaaggatc tggtaataga cttcaaaatg ttaagtgaat ttttttttct ctctatttat 2134
186 caatgatata tttcactttt aaaggaaatt ttagaggaaa attaatagct gctttttgca 2194
188 ctaaaaaacc ttgtgggtgg aaatattcct ctgagaatgg cttttatagg tattttgcct 2254
190 ggtaatgtat tcattcatga ttgcccatat tcttgaatgt ttcttcattc caatgggggtc 2314
192 aggtcaatat tatgaaaata atttttatat ttatatattgt aactaagaat ttatttctcc 2374
194 ctttactaca cgatgtaaat tcacgtcaaa ttcatgatgc tgaggattta aattcacaaa 2434
196 acctgccact acattctggt ttacattagt tacttcatgc tggctggggg tagtgaccat 2494
198 ttgcatactc ttttaaatca aggaggctgt agtagaggca gttttaagat tcttgaaggc 2554
200 aaaatttgaa aaacagtga tacttctaatt tgtttccttt tagtgccaga actaagacat 2614
202 tgtgaagcac ttgttagtaa acttaacctt gaaatgtcag actggaagga gtttttatgt 2674
204 ctttgtgcat acttctgggt attacagaaa cagtctgtaa ataacatttt aagatgcaaa 2734
206 ttttaattctg ttcacagctg atttatactg atttt 2768

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/525,020

DATE: 03/04/2005

TIME: 12:50:18

Input Set : A:\PH-1716 Sequence.TXT

Output Set: N:\CRF4\03042005\J525020.raw

```

208 <210> SEQ ID NO: 2
209 <211> LENGTH: 510
210 <212> TYPE: PRT
211 <213> ORGANISM: Homo sapiens
213 <400> SEQUENCE: 2
214 Met Ala Gly Ser Arg Gln Arg Gly Leu Arg Ala Arg Val Arg Pro Leu
215      1          5          10          15
217 Phe Cys Ala Leu Leu Leu Ser Leu Gly Arg Phe Val Arg Gly Asp Gly
218      20          25          30
220 Val Gly Gly Asp Pro Ala Val Ala Leu Pro His Arg Arg Phe Glu Tyr
221      35          40          45
223 Lys Tyr Ser Phe Lys Gly Pro His Leu Val Gln Ser Asp Gly Thr Val
224      50          55          60
226 Pro Phe Trp Ala His Ala Gly Asn Ala Ile Pro Ser Ser Asp Gln Ile
227      65          70          75          80
229 Arg Val Ala Pro Ser Leu Lys Ser Gln Arg Gly Ser Val Trp Thr Lys
230      85          90          95
232 Thr Lys Ala Ala Phe Glu Asn Trp Glu Val Glu Val Thr Phe Arg Val
233      100         105         110
235 Thr Gly Arg Gly Arg Ile Gly Ala Asp Gly Leu Ala Ile Trp Tyr Ala
236      115         120         125
238 Glu Asn Gln Gly Leu Glu Gly Pro Val Phe Gly Ser Ala Asp Leu Trp
239      130         135         140
241 Asn Gly Val Gly Ile Phe Phe Asp Thr Phe Asp Asn Asp Gly Lys Lys
242 145          150          155          160
244 Asn Asn Pro Ala Ile Val Ile Ile Gly Asn Asn Gly Gln Ile His Tyr
245      165         170         175
247 Asp His Gln Asn Asp Gly Ala Ser Gln Ala Leu Ala Ser Cys Gln Arg
248      180         185         190
250 Asp Phe Arg Asn Lys Pro Tyr Pro Val Arg Ala Lys Ile Thr Tyr Tyr
251      195         200         205
253 Gln Asn Thr Leu Thr Val Met Ile Asn Asn Gly Phe Thr Pro Asp Lys
254      210         215         220
256 Asn Asp Tyr Glu Phe Cys Ala Lys Val Glu Asn Met Ile Ile Pro Ala
257 225          230          235          240
259 Gln Gly His Phe Gly Ile Ser Ala Ala Thr Gly Gly Leu Ala Asp Asp
260      245         250         255
262 His Asp Val Leu Ser Phe Leu Thr Phe Gln Leu Thr Glu Pro Gly Lys
263      260         265         270
265 Glu Pro Pro Thr Pro Asp Lys Glu Ile Ser Glu Lys Glu Lys Glu Lys
266      275         280         285
268 Tyr Gln Glu Glu Phe Glu His Phe Gln Gln Glu Leu Asp Lys Lys Lys
269      290         295         300
271 Glu Glu Phe Gln Lys Gly His Pro Asp Leu Gln Gly Gln Pro Ala Glu
272 305          310          315          320
274 Glu Ile Phe Glu Ser Val Gly Asp Arg Glu Leu Arg Gln Val Phe Glu
275      325         330         335
277 Gly Gln Asn Arg Ile His Leu Glu Ile Lys Gln Leu Asn Arg Gln Leu
278      340         345         350

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/525,020

DATE: 03/04/2005

TIME: 12:50:18

Input Set : A:\PH-1716 Sequence.TXT

Output Set: N:\CRF4\03042005\J525020.raw

```

280 Asp Met Ile Leu Asp Glu Gln Arg Arg Tyr Val Ser Ser Leu Thr Glu
281          355          360          365
283 Glu Ile Ser Lys Arg Gly Ala Gly Met Pro Gly Gln His Gly Gln Ile
284          370          375          380
286 Thr Gln Gln Glu Leu Asp Thr Val Val Lys Thr Gln His Glu Ile Leu
287 385          390          395          400
289 Arg Gln Val Asn Glu Met Lys Asn Ser Met Ser Glu Thr Val Arg Leu
290          405          410          415
292 Val Ser Gly Met Gln His Pro Gly Ser Ala Gly Gly Val Tyr Glu Thr
293          420          425          430
295 Thr Gln His Phe Ile Asp Ile Lys Glu His Leu His Ile Val Lys Arg
296          435          440          445
298 Asp Ile Asp Asn Leu Val Gln Arg Asn Met Pro Ser Asn Glu Lys Pro
299          450          455          460
301 Lys Cys Pro Glu Leu Pro Pro Phe Pro Ser Cys Leu Ser Thr Val His
302 465          470          475          480
304 Phe Ile Ile Phe Val Val Val Gln Thr Val Leu Phe Ile Gly Tyr Ile
305          485          490          495
307 Met Tyr Arg Ser Gln Gln Glu Ala Ala Ala Lys Lys Phe Phe
308          500          505          510
310 <210> SEQ ID NO: 3
311 <211> LENGTH: 1407
312 <212> TYPE: DNA
313 <213> ORGANISM: Homo sapiens
315 <220> FEATURE:
316 <221> NAME/KEY: CDS
317 <222> LOCATION: (1)..(1071)
319 <220> FEATURE:
320 <221> NAME/KEY: sig_peptide
321 <222> LOCATION: (1)..(134)
323 <400> SEQUENCE: 3
324 atg gcg gcg gaa ggc tgg att tgg cgt tgg ggc tgg ggc cgg cgg tgc 48
325 Met Ala Ala Glu Gly Trp Ile Trp Arg Trp Gly Trp Gly Arg Arg Cys
326 1          5          10          15
328 ctg gga agg cct ggg ctt ctc ggc ccc ggc cct ggc ccc act aca cct 96
329 Leu Gly Arg Pro Gly Leu Leu Gly Pro Gly Pro Gly Pro Thr Thr Pro
330          20          25          30
332 ctc ttt ctt ctt ttg ttg ttg ggg tct gtg act gcg gat ata act gac 144
333 Leu Phe Leu Leu Leu Leu Leu Gly Ser Val Thr Ala Asp Ile Thr Asp
334          35          40          45
336 ggc aac agt gaa cat ctc aag cgg gag cat tcg ctc att aag ccc tac 192
337 Gly Asn Ser Glu His Leu Lys Arg Glu His Ser Leu Ile Lys Pro Tyr
338          50          55          60
340 caa ggg gtc ggt tcc agc tct atg ccc ctc tgg gac ttc cag ggc agc 240
341 Gln Gly Val Gly Ser Ser Ser Met Pro Leu Trp Asp Phe Gln Gly Ser
342 65          70          75          80
344 act atg ctc acg agc cag tac gta cgt ctg acc cct gac gag cgc agc 288
345 Thr Met Leu Thr Ser Gln Tyr Val Arg Leu Thr Pro Asp Glu Arg Ser
346          85          90          95

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/525,020

DATE: 03/04/2005
TIME: 12:50:19

Input Set : A:\PH-1716 Sequence.TXT
Output Set: N:\CRF4\03042005\J525020.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 13,14,16,17,19,20,25,26,28,29,31,32,34,35

Seq#:14; N Pos. 12,13,15,16,18,19,24,25,27,28,30,31,33,34

VARIABLE LOCATION SUMMARY

DATE: 03/04/2005

PATENT APPLICATION: US/10/525,020

TIME: 12:50:19

Input Set : A:\PH-1716 Sequence.TXT

Output Set: N:\CRF4\03042005\J525020.raw

Use of n's or Xaa's (NEW RULES):

Use of n's and/or Xaa's have been detected in the Sequence Listing.

Use of <220> to <223> is MANDATORY if n's or Xaa's are present.

in <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

Seq#:7; N Pos. 13,14,16,17,19,20,25,26,28,29,31,32,34,35

Seq#:14; N Pos. 12,13,15,16,18,19,24,25,27,28,30,31,33,34

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/525,020

DATE: 03/04/2005

TIME: 12:50:19

Input Set : A:\PH-1716 Sequence.TXT

Output Set: N:\CRF4\03042005\J525020.raw

L:11 M:270 C: Current Application Number differs, Replaced Application Number

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:538 M:258 W: Mandatory Feature missing, <221> Tag not found for SEQ ID#:7 ✓

L:538 M:258 W: Mandatory Feature missing, <222> Tag not found for SEQ ID#:7 ✓

L:538 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0

L:625 M:258 W: Mandatory Feature missing, <221> Tag not found for SEQ ID#:14 ✓

L:625 M:258 W: Mandatory Feature missing, <222> Tag not found for SEQ ID#:14 ✓

L:625 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0